DOCUMENT RESUME

ED 211 948 CS 006 456

AUTHOR Horowitz, Rosalind; And Others
TITLE Strategies for Classifying Readers: Effects on Prose Processing Findings.
PUB DATE 31 Mar 81
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cognitive Processes; Discourse Analysis; Elementary Education; Evaluation Criteria; Grade 6; Identification; Intermediate Grades; *Labeling (of Persons); *Reading Achievement; *Reading Comprehension; *Reading Diagnosis; *Reading Research; *Testing Problems

ABSTRACT

The limitations and advantages of any given measure of performance used to classify students as good or poor readers have not been investigated thoroughly. A study was conducted to determine what happens on several dependent measures of comprehension when the same students are separated into good and poor readers based on different criteria. Thirty-eight sixth grade students, 18 of whom were designated by their teachers as poor readers, participated in the study. The students were administered the Gates-MacGinitie Reading Comprehension Test and tests of speed and accuracy in word recognition for high and low frequency words. The students read and listened to passages, two designed to be easy and two to be difficult, that had been equated for difficulty. Following the reading and listening exercises, the students were given free recall tests and multiple choice tests of literal comprehension. The results revealed that 40% of the students shifted categories from good to poor readers or vice versa, depending upon which criteria were used for classification. The findings suggest that educators and researchers should be wary of relying on any one classification system. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED211948

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☑ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

STRATEGIES FOR CLASSIFYING READERS:

EFFECTS ON PROSE PROCESSING FINDINGS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Rosalind Horowitz
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Rosalind Horowitz, S. Jay Samuels, Matthew McGue

Center for Research in Human Learning
University of Minnesota
Twin Cities Campus

March 31, 1981

Paper presented at the American Educational Research Association Conference, Los Angeles, California, April 15, 1981, as part of Session 21.10 a symposium on Processing Differences Between Good and Poor Readers.

Preliminary Draft. Submitted for publication.

This research was supported by the Center for Research in Human Learning and a College of Education Research Grant, from the University of Minnesota.

The authors wish to thank Marcel Just for his comments on this research.

Present addresses: R. Horowitz, The University of Texas at San Antonio, College of Multidisciplinary Studies, Division of Education, San Antonio, Texas 78285

S. Jay Samuels, University of Minnesota, Center for Research in Human Learning, 205 Elliott Hall, Minneapolis, Minnesota 55455

M. McGue, Washington University, School of Medicine, St. Louis, Missouri 63110

C5086 K56

STRATEGIES FOR CLASSIFYING READERS: EFFECTS ON PROSE PROCESSING FINDINGS

Rosalind Horowitz, S. Jay Samuels, Matthew McGue
University of Minnesota

BACKGROUND

Reading comprehension is a multi-component task involving the need to decode printed material and to process for meaning. Coordinated execution of many subskills is required--skill with orthographic, semantic and syntactic levels of text. Accordingly, there may be numerous reasons for good or poor comprehension ranging from factors associated with processing mechanisms and cognitive control, within the reader, to factors such as text type and task, outside the reader.

Unfortunately, we have for some time now sought to understand differences among readers by treating comprehension as a unitary construct. It is only recently that we have developed models of reading that take into account the multi-dimensional nature of the reading act and the numerous subskills required to read for meaning. For example, Stanovich (1980) suggests that various subskills of reading (a) interact (e.g., use of top-down and bottom-up knowledge word recognition latency and use of context cues) and that these subskills can operate in a (b) compensatory manner, whereby if one skill is not operating it is replaced by another. Such an approach suggests the need to consider individual differences and what it is that is being measured in classroom diagnosis of reading tasks.

Kendall, Mason, and Hunter (1980) addressed the problem of types of reading measures used, and concluded that different kinds of tests may be measuring different aspects of comprehension. This study compared free recall, multiple-choice, cloze and maze tests and found that the highest comprehension scores were associated with maze and multiple-choice for elementary age subjects and

lower scores were associated with recall and cloze. They also found that despite great efforts to make all their passages of equal difficulty certain passages were consistently easier to comprehend than others. They concluded that one must be cautious when comparing and generalizing results from different investigations which use different measures of reading and different passages.

The study reported here is an outgrowth of research conducted at Minnesota to study differences among good and poor readers, using reading and listening tasks. This study investigates the complexity of classifying readers, in order to study differences, and is intended to be illustrative of potential problems one might encounter. Our study was designed to consider commonly used measures of comprehension, by teachers and researchers.

Overview

This study was designed in the following way: Thirty-eight students from the sixth grade were selected for the research. They were then given the Gates-MacGinitie Reading Comprehension Test and tests of speed and accuracy of word recognition for high and low frequency words. The students read and listened to passages which were equated for difficulty. Using multiple measures, two texts were designed to be easy and two other texts were designed to be hard. Then, following reading and listening, free recall and multiple choice tests were administered. This study was designed to investigate what happens on the dependent measures of comprehension, when the same set of students are separated into good and poor reader groups based on different criteria (e.g., a Gates-MacGinitie Comprehension Test or speed and accuracy of word recognition).

Questions posed were: Would experimental outcome differences among readers depend on a priori classification of those readers? Would classification methods used correlate with each other and with dependent measures? Might there be

differences in power associated with different classification methods? Finally, and most importantly, we were interested in assessing whether the subjects studied might switch categories depending upon the criteria for classification.

Rationale

Both teachers and researchers categorize students into groups such as good and poor readers following specific diagnostic procedures. Researchers use such categories in order to understand why there are differences among readers; teachers are interested in using this information to design curriculum and classroom instruction. Very often classification is based on a single measure of performance. Moreover, limitations and advantages of any given measure remain unknown. This study examined what happens on several commonly used dependent measures of prose processing and when the same students are categorized as good or poor readers.

METHOD

Determiniation of Good and Poor Readers

Subjects. Subjects were 38 sixth graders, enrolled in three upper-middle class elementary schools, outside of Minneapolis, Minnesota. They were tested in the spring of the school year.*

Teacher Judgment. Subjects were selected for this research on the basis of teacher judgment and teacher consultation with a reading coordinator in each school: 18 were identified as poor readers (that is, teachers reported that they were at least one year below grade level in the Houghton Mifflin Reading Series used in the classroom). While level in the Houghton Mifflin Reading Series was the major criteria for selection, teacher judgment was also based on day-to-day observation of student ability to perform classroom reading tasks.

---

*Note some researchers may use other dividing points, e.g., quartiles or thirds. We recognize this might affect the outcomes described here.

Using this pool of subjects, the Gates-MacGinitie Reading Comprehension and Vocabulary Test, Level D. was administered to all subjects. Results were translated into grade equivalents. For purposes of this research project, only the comprehension section of the test was used. The comprehension test includes short passages and multiple-choice questions to assess reading performance. Subjects who performed at least one year above or one year below grade level on the test were identified as good or poor respectively. Good readers had a mean comprehension grade score of 9.83 (S.D. = 1.95) and poor readers had a mean comprehension grade equivalent score of 4.26 (S.D. = 1.20).

Experimenter-Designed Tests of Speed of Word Recognition. Tests of speed of word recognition were designed by the experimenters. Speed of word recognition is considered to be an indicator of decoding automaticity.

Middle Frequency Words. Testing included 20 middle frequency words with a mean SFI Index of 49.3 (Carroll, 1971). Subjects were divided into good and poor reader groups using a median split.* The mean speed of recognition middle frequency words for good readers was 14.62 seconds and the poor readers was 27.47 seconds.

High Frequency Words. Twenty high frequency words with a mean SFI Index of 67.9 (Carroll, 1971) were used in the research. Subjects were divided into good and poor reader groups using a median split. The mean speed of recognition on high frequency words for good readers was 9.11 seconds and for poor readers was 13.11 seconds. Table I identifies diagnostic categories for good and poor readers separated on these different measures.

---

*Originally, there were 40 subjects, however, two were dropped from the study due to their absence.

Rationale for Classification Measures. The methods for classifying readers were selected for this research on the basis of their common use by teachers and researchers; they are tests that are easy to design or administer. For example, teachers frequently use a standardized test like the Gates-MacGinitie in order to assign students to reading groups. The tests of accuracy of speed of word recognition were used to classify readers into good and poor reading groups because of recent recognition that sutomaticity is an important factor in higher-level text processing and in comprehension.

Reading and listening tasks were included because wheras automaticity in decoding appears to be important in reading, it should be of little importance in listening comprehension. Thus, including these two tasks (reading and listening) allows one to determine if a comprehension deficit is due to decoding problems in reading, limitation in use of prosodic features of text, or represents some other processing problem, e.g., a more general comprehension problem characteristic of both reading and listening.

Texts. Easy texts used in this research were written at 4/5th grade readability and hard texts were written at 7/8th grade readability (Dale, Chall, 1948). Easy texts were 300 words in length and hard texts were 367 words in length. Texts were also controlled for number of idea units and subjects were tested three weeks before the actual experiment to determine their degree of prior knowledge of the topic and content. A Summary of measures used to control for text difficulty are found in Table II.

Insert Table II About Here

Scoring. Free recalls were scored for number and level of literal idea units (following Meyer, 1975). The multiple-choice test was designed to assess literal comprehension and included items extracted from all levels of the text hierarchy. The multiple-choice questions were presented orally in the listening condition and were presented on paper to be read for the reading condition. The same procedure was used for scoring free recall under listening and reading conditions.

RESULTS

Shift in Categories. We find that students classified as good or poor readers may shift categories depending upon the method used to classify them. As Table III indicates, a number of students shifted from our good to poor reader category or vice versa depending upon the classification method or diagnostic procedure used.

Insert Table III About Here

Significant Differences Among Readers. We find that using the same dependent variable one can obtain significant differences or no significant differences between good and poor reader groups, depending upon the a priori method of classification. To elaborate, there is no difference in the results obtained when one examines main effects. That is, in this study, all four classification precedures provided similar results (either significant or no significant differences) across reader groups. However, when one turns to the interactions, significance or no

significant difference is based upon the classification procedures used at the onset of the experiment. It was hypothesized that some a priori classification and diagnosing might be more correlated with select dependent measures used in the study, than other classifications.

Main Effects. There was consistently significant differences between good and poor readers across measures when the dependent variable was number of decoding errors ($p < .00$) and speed of reading in words per minute (WPM) ($p < .00$) but consistently no significant differences between these good and poor readers, on all four classification procedures, when the dependent measure was number of literal idea units recalled. The multiple-choice test resulted in significant difference and no significant difference depending upon the classification method.

Main Effects: Good vs. Poor Reader Differences

Based On A Priori Classifications

Dependent Measure	Gates-MacGinitie Comprehension Test - Level D	Automaticity Middle Frequency Words	Automaticity High Frequency Words	Teacher Judgment
Decoding Errors While Reading	$p < .001$*	$p < .005$*	$p < .007$*	$p < .001$*
Reading Speed WPM	$p < .000$*	$p < .000$*	$p < .000$*	$p < .00$*
Literal I.U.'s Recalled	$p = .197$	$p = .392$	$p = .989$	$p = .519$
Multiple Choice	$p < .000$*	$p = .07$	$p = .44$	$p < .01$*

*Significant Difference

Interactions. Ability x Text Difficulty. When comparing good and poor reader groups on easy versus hard texts, the following results were found. When classified by speed of word recognition (middle frequency words) and teacher judgment, there was a significant interaction by ability (between good and poor reader groups) on easy and hard texts--on number of literal idea units recalled ($p < .03$) on each, but no significant difference with the Gates-MacGinitie or speed of recognition of high frequency words.

Good vs. Poor A Priori Classification

Dependent Measures	Gates-MacGinitie Comprehension Test - Level D	Automaticity Middle Frequency Words	Automaticity High Frequency Words	Teacher Judgment
Literal I.U.'s	$p = .18$	$p < .03$*	$p = .58$	$p < .02$*
Multiple Choice	$p = .11$	$p = .06$	$p = .85$	$p = .84$

* indicates significant difference

Ability x Mode. When examining good and poor reader groups on listening or reading, there was a significant interaction ($p < .02$) with the Gates-MacGinitie Classification method on number of literal idea units recalled but not with the other measures for classification.

Ability x Difficulty x Mode. A three-way interaction was significant with the Gates-MacGinitie classification ($p < .03$) and teacher judgment ($p < .03$) on literal idea units recalled but not with other measures.

Interpretation of Findings

In order to futher assess these outcomes, we chose to examine the pattern of one of the interactions, Ability x Text Difficulty, using literal idea units recalled as the dependent measure.

Using this as a prototype raised the following questions? a) Did the same pattern manifest itself; b) Did the amount of experimental error vary as a function of classification scheme; and c) Were the various measures of reading performance highly correlated?

Insert Figure 1 About Here

As Figure 1 indicates, there is not a substantive difference among the interactions. The patterns are much the same.

Note that significant differences among readers x text emerge with speed of recognition of middle frequency words and with teacher judgment--but not the Gates-MacGinitie or high frequency words.

Further analysis of the data revealed that the error variances associated with each classification procedure were not the same. As the list below indicates, there is less error associated with the automaticity recognition of middle frequency words and teacher judgment--than there is with the other criteria. It should be noted that, in this instance, the greatest amount of error was associated with the Gates-MacGinitie test, a test commonly used by teachers and researchers to partition populations.

ERROR VARIANCES FOR EACH OF THE
CLASSIFICATION PROCEDURES USING LITERAL IDEA UNITS
RECALLED AS THE DEPENDENT MEASURE

Gates-MacGinitie	315.31
Automaticity High Frequency Words	312.28
Automaticity Middle Frequency Words	288.21
Teacher Judgment	287.02

Additionally, correlations were run to determine the correlation of each classification method with the dependent measure literal idea units recalled. We find that the greatest correlation between the Gates-MacGinitie and the literal idea units recalled, .55. Note that automaticity with middle frequency words is correlated .49 with the dependent measure about the same degree as teacher judgment, .45.

CLASSIFICATION METHODS AND CORRELATIONS WITH LITERAL IDEA UNITS RECALLED
(summing easy, listen, easy read plus hard listen, hard read)

Gates-MacGinitie	.55
Middle Frequency Words	.49
High Frequency Words	.26
Teacher Judgment	.45

Finally, we also investigated correlations among the four measures of reading. Note that the correlation between teacher judgment and automaticity middle frequency words was .74.

Insert Table IV About Here

Conclusions

These results suggest that the same subjects may appear in good or poor reader categories (high or low ability groups) depending upon which tests are used to classify them. In this study, 40% of our students shifted categories depending upon which criteria were used for classification. Table III identifies those students who shifted from our good to poor reader category or vice versa.

Table V notes the number of students who were judged differently when comparing them on two measures of reading performance.

The fact that the same student may shift category from good to poor reader has implications for educators. For example, students may be classified as having a reading problem under one set of conditions but not another. Also, researchers who assign students to categories should be aware of possible unreliability of some classifications and that the unreliability of classifications may affect the validity of findings.

References

Carroll, et al. Word Frequency Book, Boston: Houghton Mifflin, 1971.

Daneman, M. and Carpenter, P. Individual Differences in Working Memory and Reading. Journal of Verbal Learning and Verbal Behavior, 19, 450-66 (1980).

Hildyard, A. The emergence of literate competence. Paper presented at AERA, San Francisco, 1979.

Hunt, E. We know who knows, but why? In Schooling and the Acquisition of Knowledge, (Editors, Anderson, Spiro, Montague). Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Kendall, J. R., Mason, J., Hunter, W. Which Comprehension? Artifacts in the Measurement of Reading Comprehension. The Journal of Educational Research. March, April 1980, Volume 73, Number 4, 233-36.

Kintsch, W., and Kozminsky, E. Summarizing stories after reading and listening. Journal of Educational Psychology, 69, 5, 1977, 491-99.

Meyer, B. The organization of prose and its effects on memory. Amsterdam: North-Holland, 1975.

Mosenthal, P. Children's comprehension of different structures in two modes of processing. Mimeo. SUNY, Albany, n.d.

Perfetti, C. A. and Lesgold, A. M. Discourse comprehension and sources of individual differences in M. A. Just and P. A. Carpenter (Eds.) Cognitive processes in comprehension. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Perfetti, C and Hogaboam, T. Relationship between single word decoding and reading comprehension skill. The Journal of Educational Psychology, 1975, Volume 67, Number 4, 461-9.

Rubin, A. A theoretical taxonomy of the differences between oral and written language. Center for the Study of Reading, Technical Report No. 35, January, 1978.

Stanovich, K. Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, Number 1, 1980.

Sticht, T. Auding and reading: a developmental model. Alexandria, Virginia, Human Resources Research Organization, 1974.

TABLE I

DIAGNOSTIC CATEGORIES FOR GOOD AND POOR READERS
SEPARATED ON DIFFERENT MEASURES

Gates-MacGinitie Reading Comprehension Test Level D.

(Reported in Grade Equivalents)

Good	Poor
Mean = 9.83	Mean = 4.26
(S.D. = 1.95)	(S.D. = 1.2)

Middle-Frequency Words -- Speed of Word Recognition

(for 20 words reported in seconds)

Good	Poor
Mean = 14.62	Mean = 27.47
(S.D. = 4.54)	(S.D. = 13.86)

Number Words Correct out of 20

Mean = 19.1	Mean = 14.79
(S.D. = 1.7)	(S.D. = 3.06)

High-Frequency Words -- Speed of Word Recognition

(for 20 words reported in seconds)

Good	Poor
Mean = 9.11	Mean = 13.11
(S.D. = 2.32)	(S.D. = 9.29)

Number Words Correct out of 20

Mean = 20.0	Mean = 19.00
(S.D. = 0)	(S.D. = 2.51)

TABLE II

MEASURES OF TEXT DIFFICULTY

	EASY		HARD	
	Text I	Text II	Text III	Text IV
Dale-Chall Formulae	4th/5th	4th/5th	7th/8th	7th/8th
Number Words	300	300	367	367
Number Idea Units	106	130	137	110
Literal Idea Units Recalled	$\bar{x}$=43.05	$\bar{x}$=37.58	$\bar{x}$=21.90	$\bar{x}$=16.83
(Percent of Total)	33.15%	35.00%	15.98%	14.77%
Prior Knowledge on Topic (in idea units)	$\bar{x}$=26.43	$\bar{x}$=26.48	$\bar{x}$=20.98	$\bar{x}$=22.25
Topic	Eskimos	Seagulls	Earthquakes	Date Trees

Strategies for Classifying Readers

Sixth Graders	Good 1	Good 2	Good 3	Good 4	Good Total	Proportion Good
		Indicators of Decoding Automiticity				
Subject Number	Gates-MacGinitie Comprehension Test Level D	Middle-Frequency Words ($\bar{X}$ SFI Index) (49.3)	High-Frequency Words ($\bar{X}$ SFI Index) (67.9)	Teacher Judgment		
1	1.	0	0	0	1.	.25
2	0	0	0	0	0	0
3	1.	1.	1.	1.	4.	1.00
4	1.	1.	1.	1.	4.	1.00
5	0	0	0	0		0
6	1.	1.	1.	1.	4.	1.00
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	1.	1.	0	1.	3.	.75
10	0	1.	1.	1.	3.	.75
11	0	0	1.	0	1.	.25
12	0	0	0	0	0	0
13	1.	1.	1.	1.	4.	1.00
14	1.	0	0	0	1.	.25
15	1.	0	0	0	1.	.25
16	1.	1.	1.	1.	4.	1.00
17	1.	1.	0	0	2.	.50
18	1.	1.	1.	1.	4.	1.00
19	0	1.	0	0	1.	.25
20	1.	0	1.	0	2.	.50
21	0	0	0	0	0	.0
22	0	0	1.	0	1.	.25
23	1.	1.	1.	1.	4.	1.00
24	1.	1.	1.	1.	4.	1.00
25	1.	0	1.	1.	3.	.75
26	0	0	0	0	0	0
27	1.	1.	0	1.	3.	.75
28	1.	1.	1.	1.	4.	1.00
29	0	0	0	0	0	0
30	0	0	1.	0	1.	.25
31	0	0	0	0	0	0
32	1.	1.	1.	1.	4.	1.00
33	1.	1.	1.	1.	4.	1.00
34	1.	1.	1.	1.	4.	1.00
35	0	1.	1.	0	2.	.50
36	0	0	0	1	1.	.25
37	1.	1.	1.	1.	4.	1.00
38	0	0	0	0	0	0
39	0	0	0	0	0	0
40	0	1.	0	0	1.	.25

Key: Good = 1.
Poor = 0

Carroll, et al. Word Frequency Book. Boston: Houghton-Mifflin, 1971, was used to obtain $\bar{X}$ SFI Indices.

TABLE III

TABLE V

Gates-MacGinitie Comprehension Test Level D

Middle Frequency Words

		Poor	Good	
Teacher Judgment	Good	2	16	18
	Poor	17	5	22
		19	21	40

High Frequency Words

		Poor	Good	
Teacher Judgment	Good	2	16	18
	Poor	18	4	22
		20	20	40

		Poor	Good	
Teacher Judgment	Good	4	15	19
	Poor	16	5	21
		20	20	40

Middle Frequency Words

		Poor	Good	
Gates-MacGinitie Comprehension Level D	Good	5	16	21
	Poor	15	4	19
		20	20	40

High Frequency Words

		Poor	Good	
Gates-MacGinitie Comprehension Level D	Good	6	15	21
	Poor	14	5	19
		20	20	40

High Frequency Words

		Poor	Good	
Middle Frequency Words	Good	5	15	20
	Poor	15	5	20
		20	20	40

FIGURE 1

ABILITY x TEXT DIFFICULTY

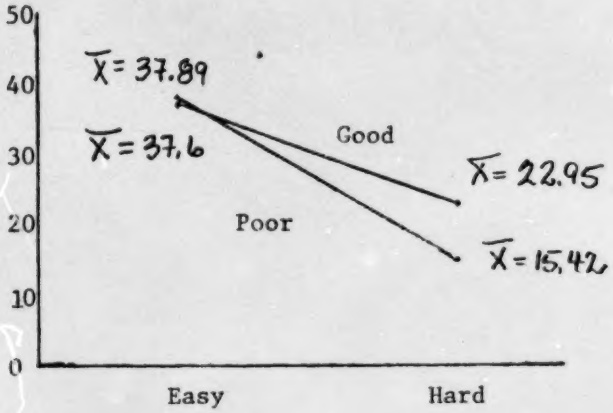
LITERAL IDEA UNITS

(Combines comprehension of literal idea units on read and listen conditions)

Interactions

Gates-MacGintie Reading Comprehension Test Level D

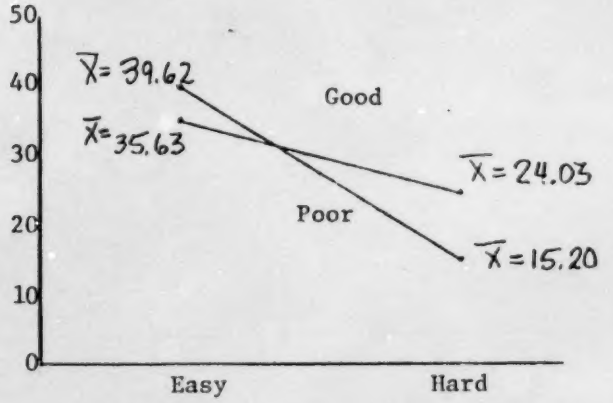
Literal Idea Units



p=.20

Speed of Word Recognition

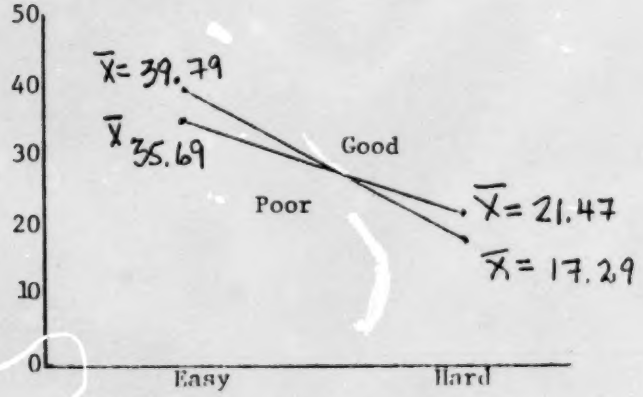
Middle Frequency Words



p<.03 *

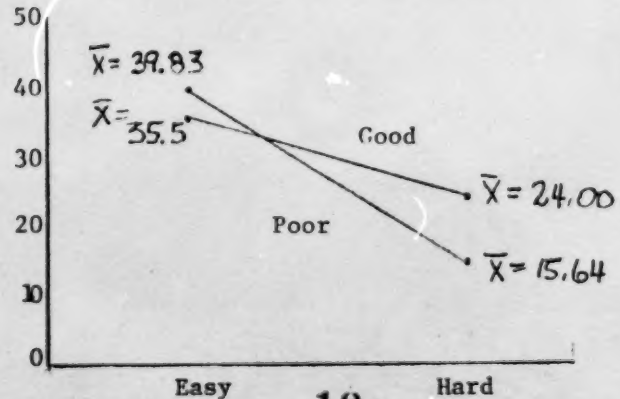
Speed of Word Recognition

High Frequency Words



p=.16

Teacher Judgement



p<.02 *

* Significant Difference